Algorithms for the Feature Extraction of Electro Cardiograms for use in Implantable Cardioverter Defibrillators :A Review

Gopika P
Sree Chitra Tirunal Institute
For Medical Sciences and Technology
Thiruvananthapuram, India.

Jithin Krishnan
Sree Chitra Tirunal Institute
For Medical Sciences and Technology
Thiruvananthapuram, India.

Muraleedharan C V
Sree Chitra Tirunal Institute
For Medical Sciences and Technology
Thiruvananthapuram, India.

***Abstract*—Fora cardioverter defibrillator, the live electrocardiogram wave feature extraction for analyzing and classifying of arrhythmias forms the vital part of algorithm. The fiducial points which form the inputs for the extraction and classifying algorithms were derived out of existing highly popular Pan-Tompkins algorithm, advanced wavelet transform based algorithms and Threshold Crossing Intervals (TCI) algorithm. The most common algorithms employed for the purpose are reviewed for their efficacy and specificity in this article**

Keywords*— Implantable Cardioverter Defibrillator, Pan Tompkins Algorithm ,Discrete Wavelet Transforms(DWT) , TCI algorithm, Morphology algorithm, Arrhythmia.)

I. INTRODUCTION

An implantable cardioverter-defibrillator (icd) is a device which is able to perform cardioversion defibrillation, and anti-tachycardia pacing of the heart. The device is therefore capable of correcting most life-threatening cardiac arrhythmias. The icd is the first-line treatment and prophylactic therapy for patients at risk for sudden cardiac death due to ventricular fibrillation(vf) and ventricular tachycardia(vt).

Commensurating with the statistics of world health organization, 17.7 million people die from cardio vascular diseases (cvd) per year. In india the agestandardized cvd is 272 per 100,000 population, which is higher than the global average of 235 per 100,000 population and this alarming situation in a developing country calls for indigenous solution to cater the needs of a fast growing population[1].

The current study features a preliminary analysis in ecg feature extraction which plays a key role in identifying the various cardiac arrhythmias which in turn is used to analyze, detect and confirm the arrhythmia condition, for which appropriate therapies needed to be decided and delivered. In this review paper, emphasis is given on selection of an appropriate algorithm for implementing in a defibrillator. Various literatures available in the field are studied first. This has helped in identifying the problem. The paper next describes the methodology adopted for arriving at a conclusion for the selection of an appropriate algorithm which helps in feature extraction of ecg signals.in the final section of the paper the results obtained were discussed.

II. ICD ALGORITHMS

Electrocardiography(ecg) feature extraction has been studied from early time and lots of advanced methods as well as transformations have been proposed for this purpose. A plethora of excellent algorithms have been developed based on different signal processing approaches such as time-domain morphology analysis, wavelet transform based analysis, statistical parameters based analysis and other techniques such as artificial neural networks, pattern matching, hidden markov models or combinations of above methods [3,4,7,8].

Zhao et al. [3] proposed a feature extraction method using wavelet transform and support vector machines. The proposed system is comprised of three components including data pre-processing, feature extraction and classification of ECG signals. Two diverse feature extraction methods are applied together to obtain the feature vector of ECG data for reliable heart rhythm recognition. Here the features of the ECG segment are derived from coefficients of the wavelet. Concurrently, autoregressive modelling (AR) is also applied to get the temporal structures of ECG waveforms. This method is found to achievethe overall accuracy of 99.68%

Jiapu Pan and Willis J. Tompkins [6] developed a real-time algorithm for detection of the QRS complexes of ECG signals. In this algorithm, the QRS features are recognized based upon digital analyses of slope, amplitude, and width. A digital band pass filter is designed to reduce false detections caused by the various types of noises and interference present in ECG signals. This filtering permits use of low thresholds, thereby increasing detection sensitivity. The algorithm provides an adaptive threshold and to compensate for ECG changes as QRS morphology and heart rate. Thismethod is found to achieve the overall accuracy of 99.3% for the standard arrhythmia database waveforms.

Evangelos, Dwaipayan Biswas, et al[7] developed a low-complexity algorithm for the extraction of the fiducial points from the ECG. The algorithm employs DWT with the Haar function as the mother wavelet. From the maximum and minimum point's analysis on the DWT coefficients, an approximation of the ECG fiducial points is extracted. This is later added with a refinement stage which improves the temporal resolution of the DWT.

Nitish V Thakor, et al [14] developed a sequential hypothesis testing algorithm which is able to detect the arrhythmias from the statistical parameters of the ECG signal. Here a TCI is calculated and a probability distribution is plotted based on this TCI. The probability distribution provides the means for successful classification of arrhythmias. This algorithm needs to be tested on a larger database.

Thakor et al[15] developed a process which utilises the nonlinearity of Sinus Rhythm(SR), VT and VF signal. Here the given ECG signal is converted to a binary data based on a calculated threshold. The signal is further classified based on the probability distribution calculated of average TCI of SR,VF and VT.

Brown et al [16] has developed a wavelet-based morphology algorithm which help discrimination of shockable and non-shockable arrhythmias.

III. IMPLEMENTATION

The strategy was to study existing ECG feature extraction algorithms to extract the ECG fiducial points that are clinically important. The following four popular primary algorithms were implemented:

1. Pan Tompkins Algorithm
2. Time domain based method of TCI algorithm
3. DWT-MMA Algorithm
4. Wavelet based method of morphology algorithm.

IV. METHODOLOGY

A. *Electrocardiography waveform study*

ECG is the process of recording the electrical activity of the heart over a period of time using electrodes placed on the skin. These electrodes detect the tiny electrical changes on the skin that arise from the heart muscle's electrophysiological pattern of depolarizing and repolarizing during each heartbeat.These waveforms are labeled P, Q, R, S and T [5]. P wave is the first short upward movement of the ECG tracing. It indicates that the atria are contracting, pumping blood into the ventricles. The QRS complex normally beginning with a downward deflection (Q), larger upwards deflection R and then a downwards S wave. The QRS complex represents ventricular depolarization and contraction. The PR interval indicates the transit time for the electrical signal to travel from the sinus node to the ventricles. T wave is normally a modest upwards waveform representing ventricular repolarization. This feature extraction scheme mainly determines the amplitudes and intervals in the ECG signal which determines the functioning of heart of every human. A Basic ECG waveform showing PQRST segments is shown in Fig. 1.

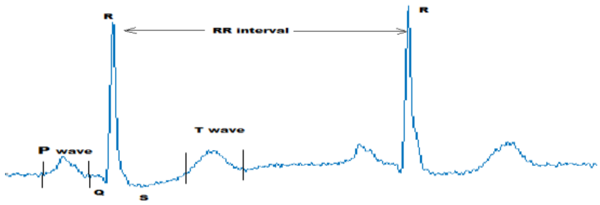


Fig.1 Basic ECG waveform showing PQRST

B. *Need for ECG Analysis*

The ECG Feature Extraction plays a vital role in diagnosing most of the cardiac disease. Therefore, development of accurate algorithms for automatic ECG feature extraction is of major importance and also the system should perform accurately. In feature extraction, the amplitude and width of the basic waves and segments within ECG signal are detected so as to allow successful abnormality detection. Usually ECG feature extraction is carried out through three steps such as pre-processing, feature detection and heart beat classification. In pre-processing the various noises present in the ECG is filtered by selecting the appropriate filters. In the next stage fiducial points and peaks are detected and using this information the heart rhythm is classified in the last stage.

C. *MATLAB Implementation-Initial Study*

The ECG samples under study were taken from Physionet [17]data base. The ECG waveform selected is of one minute duration and has a sampling rate of 125 samples/sec with a total of 7500 samples points. The process involves three steps:

1) High pass filtering
2) Thresholding
3) Segmentation.

After the high pass filtering with a cut off frequency of 40Hz using and Direct Form II transposed filter, thresholding is applied with a threshold value of 0.5 mV to detect R peaks. The heart rate is calculated by measuring the number of R peaks. Then using R peak as reference, segmentation is done to find P, Q, R,S and T waves. The final labelled ECG wave is shown in Fig. 2.

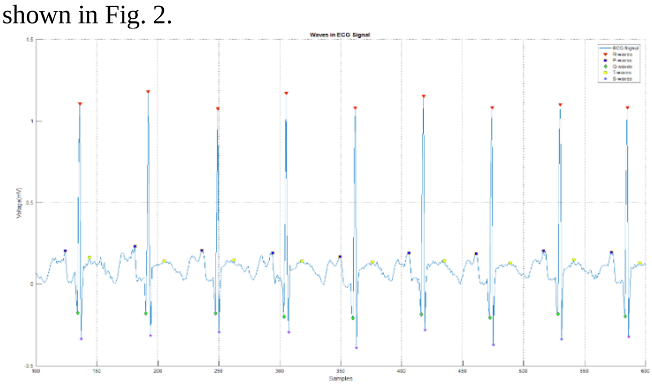


Fig 2 Initial detection of PQRST

The problem with this method is it uses a fixed threshold to identify R wave, rather than an adaptive thresholding which adjusts to the disparities of signal due to various noises.

Study of Arrhythmias

The term arrhythmia (also called dysrhythmia) is very general, referring to all rhythms other than the normal rhythm of the heart, sinusrhythm(sr).normal sinus rhythm is regular with a heart rate between 60 and 100 beats per minute. The p waves are normal in size, shape and direction.

The most frequent cause of suddencardiacdeath is ventricular arrhythmias i.e.vf and vt.vt is an arrhythmia in which heart rate increases to the order of 140 to 250 beats per minute. The rhythm is usually regular, but may be slightly irregular. Vt occurs as a series of wide qrs complexes seen in short runs or as a continuous rhythm. In vf a disorganized, chaotic, electrical focus in the ventricles take over control of the heart. There will be no qrs complex, but instead the ventricular muscle quivers and is often described as 'resembling a bag of worms'.fig.3 shows a way to compare the rhythms based on their heart rate, pr interval, rr interval and qrs width.

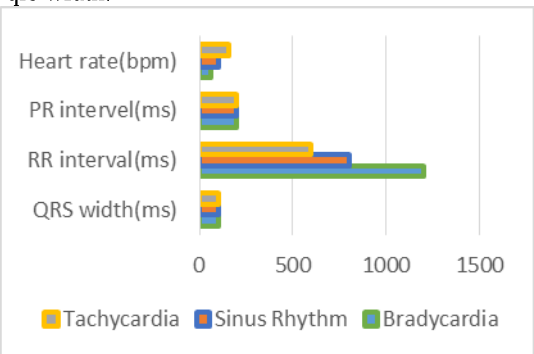


Fig 3 Arrhythmia identifying parameters

To study the arrhythmias a GUI in MATLAB (R2016a) was developed [6].Fig.4 shows a running display of the arrhythmias which the user can select from a dropdown list.

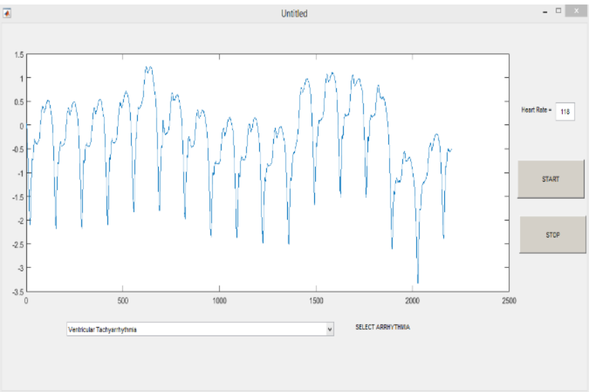


Fig 4 GUI for arrhythmia study

Lightweight waveform and annotation viewer and editor is a provision in Physionet data base used to view any of the recordings of physiologic signals and time series in the database, together with their annotations (event markers) prior to downloading and viewing in Matlab.

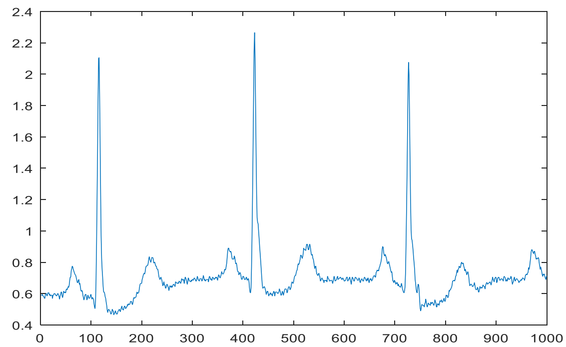
D. Study of Algorithms

1) PAN TOMPKINS ALGORITHM

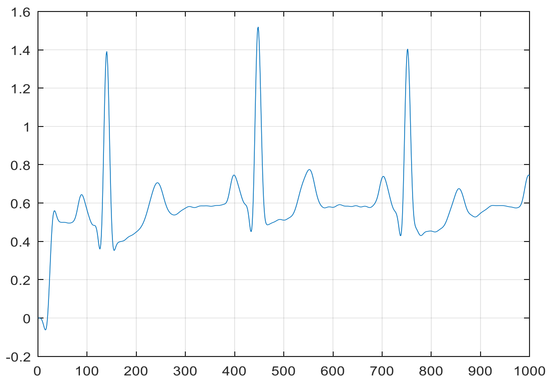
This is a time-domain based morphology analyzing algorithm which detects QRS complex in ECG signals[3]. The QRS complex represents the ventricular depolarization of the heart.

The main processes involve band pass filtering Fig. 5(b), differentiation (Fig.5c), squaring (Fig.5d) and moving window integration (Fig 5 e). The thresholds applied here are adaptive and the coefficient values were given in the algorithm.

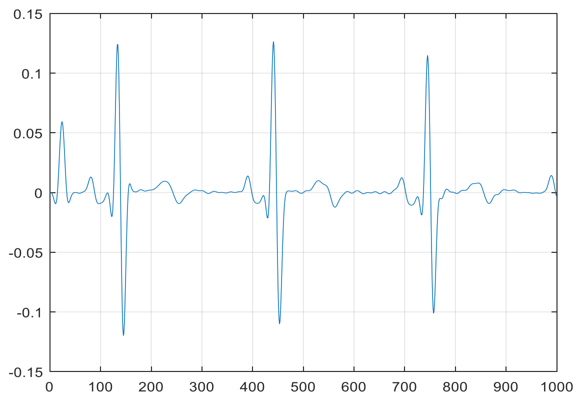
a)



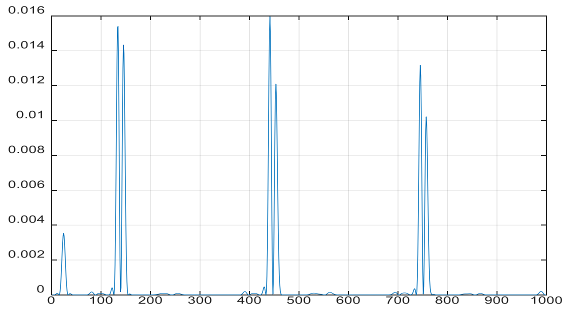
b)



c)



d)



e)

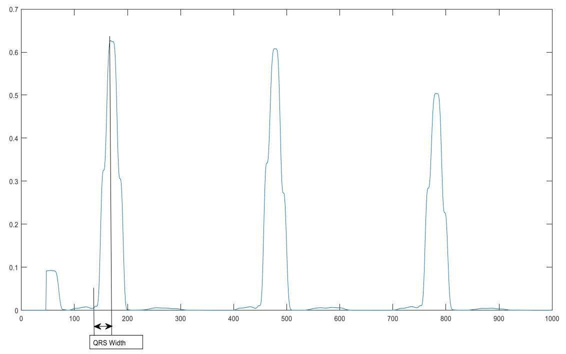


Fig.5. Pan-Tompkins algorithm a) Original signal b)Band pass filtered output c)Differential output d)Squared waveform e)Output of moving window integration. Its half width is equivalent to QRS width.

A band-pass filter of order 48 and cut-off frequencies of 3.6Hz and 36 Hz was designed.The filter process delay was found to be 30 samples. For the derivative function a five-point derivative was used in the original algorithm. Here a two-point derivative is used with the delay of 1 sample.

The squaring makes all data-points positive, amplifies the output and also restricts false positive. The moving window integration was done using window length of 15 samples. The window length was selected after several iterations. The adaptive thresholding uses the same equation and coefficients used in the original algorithm and gave the best result. Heart rate is calculated using the formula(1)

$$\frac{60 * fs}{RRwidth} \qquad (1)$$

The fiducial points derived using this algorithm are QRS width and R peak location. QRS complex corresponds to the rising edge of the moving integration waveform and the peak denotes the R peak.

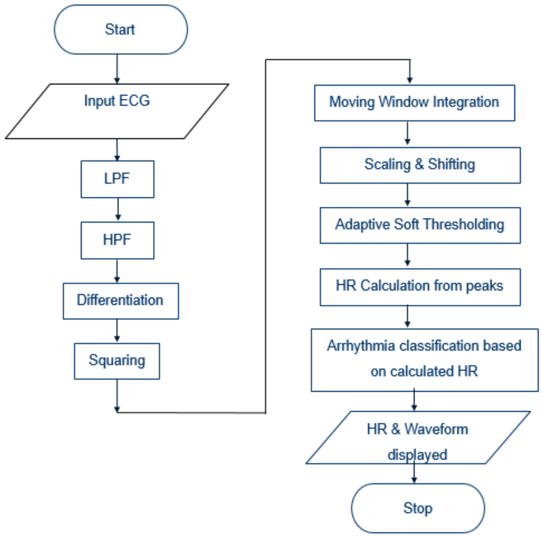


Fig.6. Pan-Tompkins algorithm

The efficiency of the algorithm was calculated using the formula(2).

$$Efficiency = \frac{OHR - |OHR - DHR|}{OHR} * 100 \qquad (2)$$

where OHR = Original Heart Rate and DHR = Detected Heart Rate. The efficiency was found to be 98.5% when tested with ECG record No. 101 in MIT-BIH database. Performance of the algorithm was evaluated and is shown in Table 1 In the table

TP = True Positive
FP = False Positive
TN = True Negative
FN = False Negative

Further analysis of the program was carried out on the basis of RR interval and QRS width obtained for different ECG signals as seen in Fig.7.

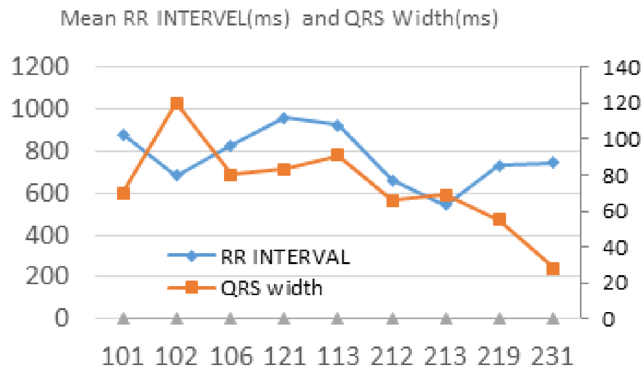


Fig.7.Mean RR interval and QRS width

TABLE I. FEATURE EXTRACTION PERFORMANCE FROM MIT DATABASE

Data record	R detection
101	TP
102	FP
106	TP
121	TP
113	TP
200	FN
212	TP
213	TP

All the signals tested above are Sinus Rhythm.All the above data signals were chosen from MIT-BIH arrhythmia database. This database contains half hour long excerpts from 48 subjects of Beth Israel Hospital which supports worldwide research on arrhythmia analysis.They are sampled at 360 samples/sec and has 11-bit resolution in range of 10mV.Physionet have been providing this database at free of cost.

Upon running the same algorithm with Ventricular Tachycardia the QRS width was not obtained correctly.This algorithm took 290 millisecond to evaluate 1sec of data (360 samples).

2) DISCRETE WAVELET TRANSFORM-MAXIMUM MODULUS ANALYSIS(DWT-MMA) ALGORITHM

Another line of approach for ECG waves feature extraction is based on wavelet transform [9]. The advantage of using wavelets is that frequency domain filtering is performed implicitly during DWT computation.

Periodic occurrence of patterns with different frequency contents will make an ECG signal. Therefore, Wavelet Transforms are an excellent tool for its analysis. This method is a correlation between the given signal and the mother wavelet such as Haar, Daubechies, Coiflets ,Symlets etc.

The output of each level of wavelet will give approximate and detailed coefficients. Wherever a high value is obtained from these coefficients we consider them as significant points and the lower valued insignificant points are found to have low correlation with the mother wavelet chosen. The aim is to discard those points which does not contribute much to the information of the signal so as to reduce computational complexity. It is evident that most of the energy of the ECG signal lies within the levels of 2^1 to 2^5. Thus higher levels of wavelets were not considered in this analysis.

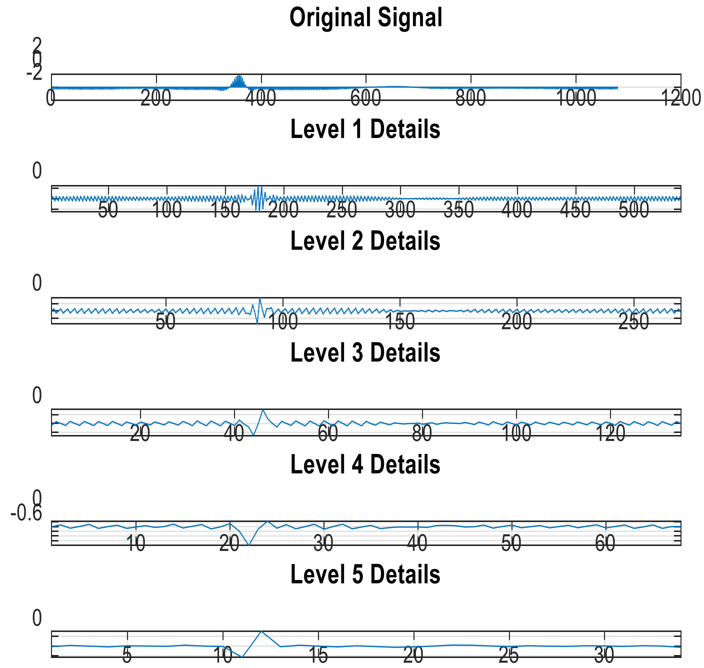


Fig. 8. Detailed coefficients of a PQRST complex

The criterion chosen in this work is to develop a low complexity algorithm for automatic ECG analysis.

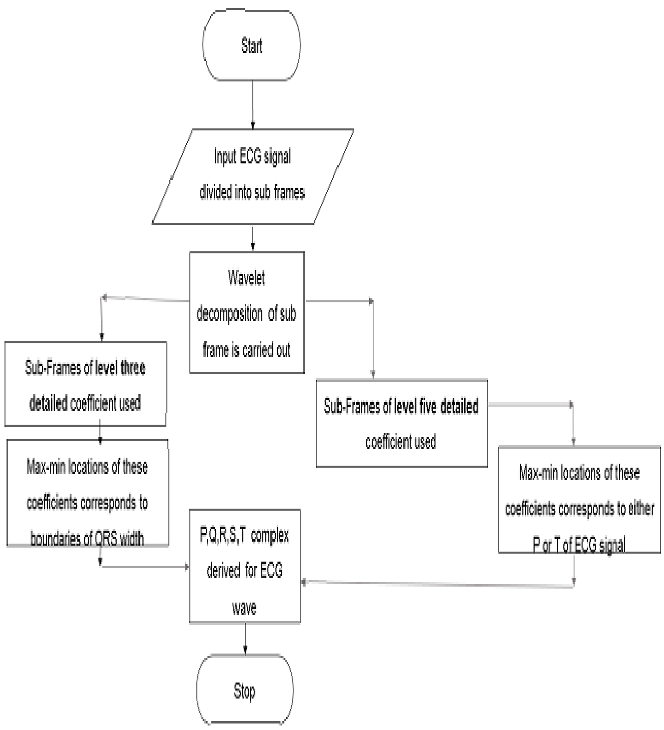


Fig. 9. DWT – MMA algorithm

The algorithm starts by taking a PQRST complex. The mother wavelet chosen for DWT is critical and the most simplest wavelet 'Haar' was considered.

From Fig. 8. it is obvious that the first two levels of detailed coefficients contains noise components. From level three onwards, the signal is devoid of noise and hence level three detailed coefficients are considered for QRS complex and level 5 detailed coefficients were used for P and T waves.

First step in the algorithm utilizes the 2^3 detailed coefficients to get the QRS complex. The maximum and minimum locations of this coefficients are taken which identifies the boundaries of QRS complex in the corresponding ECG signal. The peak among this sub-array is taken as the location of R peak.To get the P wave, the portion of the wave to the left of this QRS complex is taken and again the maximum and minimum were found. This sub-array corresponds to P wave provided that it satisfies the clinical criterion of a P wave. Again maximum-minimum pair to the right of the QRS complex is analyzed to find the T wave, if it fits into the medical standard of a T wave.

The algorithm was tested using different ECG signals. The result shown below is of Record '103'of MIT-BIH database which contains a Sinus Rhythm ECG signal.

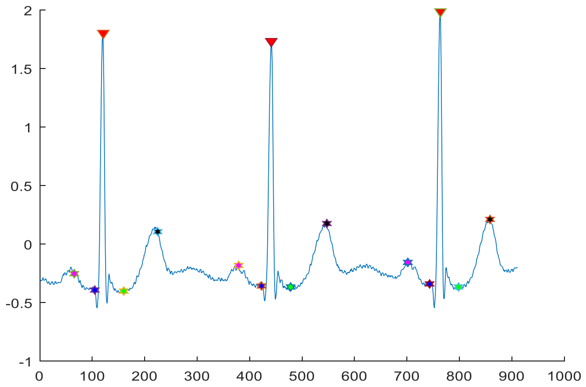


Fig. 10. PQRST complex identified through DWT – MMA algorithm

The computational complexity of the algorithm is considered next using signals of length 'N'. It requires N/2 additions and N/2 multiplications for cD-l1 coefficients. Similarly, cD-l2 requires N/4 additions and N/4 multiplications. In the DWT-MMA method only cD-l3 and cD-l5 are used. The method based on Mallat's algorithm requires N/2 + N/4 + 2N/8 + N/16 + N/32 additions and multiplications each.

TABLEII Performance Analysis of Dwt-Mma Algorithm

Record Name	**RR interval(ms)**	**QRS Width(ms)**
101	850	130
102	808	220
106	540	130
200	910	130
121	970	150
113	540	130

This algorithm will require further modification for identifying VT signals. It took 600 millisecond for the algorithm to give the result for 1 second of data.

3)TCI ALGORITHM

This algorithm works on the statistical parameters of the data signal to be tested and is purely time domain based. It works on the basis of a Threshold Crossing Interval estimation. A decision is made based on this TCI calculated.

The data signal for testing is first divided into partitions of 1sec duration. For each partition the TCI is calculated based on the equation (3)

$$TCI = \frac{1000}{(N-1)+\frac{t_2}{t_1+t_2}+\frac{t_3}{t_3+t_4}} \quad (ms) \qquad (3)$$

'N' is obtained by number of pulses in the 1sec signal included between the dotted red line as indicated in Fig.13.'t1' is the duration in time required for the pulse to complete in the preceding time frame of the signal. 't2' is the duration in time required for the pulse to complete in the beginning of selected time frame of the signal. Towards the end of the selected time frame, 't3' will show the time frame of starting of the pulse. The corresponding pulse ending width is denoted by' t4'.

For each partition when the signal exceeds the TCI, a binary 1 is assigned when the signal is below the TCI a binary 0 is assigned. Thus a new binary signal is created from the original data signal.

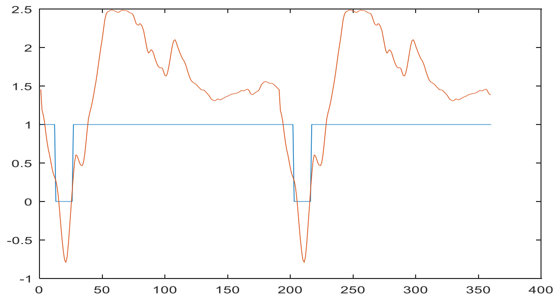


Fig. 11.A VT signal and its corresponding binary sequence

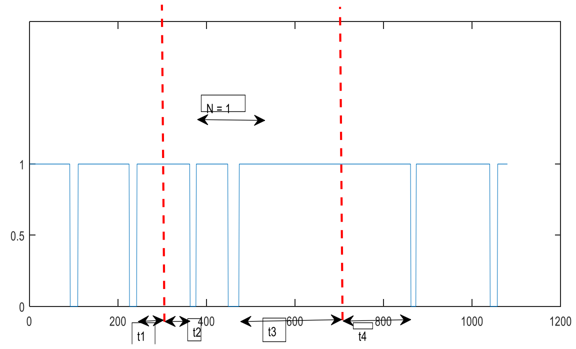


Fig.12. Obtaining N, t1,t2,t3 and t4 for TCI calculation

For the signal shown in the figure TCI obtained was 359.32. It is declared as a VT signal if 220<TCI<500[14]. Thus the tested signal was successfully classified by the TCI algorithm as VT signal.

TABLEIII CLASSIFFICATION USING TCI ALGORITHM

Signal Type	**Range of TCI(ms)**	**Mean(ms)**	**Standard Deviation**
Normal	450-850	650	13.5
VT	0-500	220	16.5
VF	0-200	105	6.5

But the disadvantage of this method is that there is a small overlapping region for VT and SR signal. The VF and VT signal will have a larger overlapping area. This creates a situation for unwanted therapies [16]. So further modification needs to be adopted before concluding for the therapy to be delivered.

4) MORPHOLOGY ALGORITHM

In this algorithm morphology of a template is created from a previous beat and compared with the morphology of the present beat to discriminate between supraventricular tachycardia from ventricular tachycardia. The templates are calculated from the wavelet transforms of the selected beats. A match percentage is calculated through correlation from these templates and classification is done based on the value obtained by other researchers [16][18][19]. From the experiments done on VT signals obtained from MIT-BIH it was observed the match percentage to be between 75 and 83 for VT signals.

This method is relatively more popular method. It is already being used in many existing ICD devices but is more complex than the time domain analysis methods. But once the decision is made this is found more reliable.

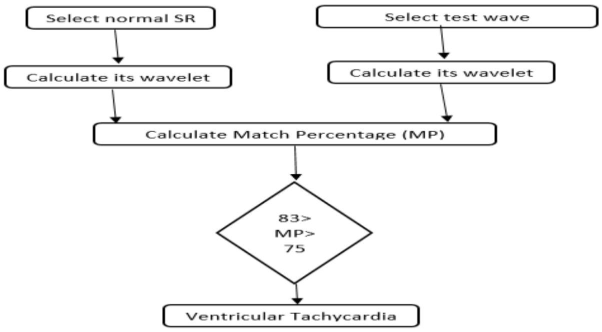


Fig. 14. Morphology algorithm

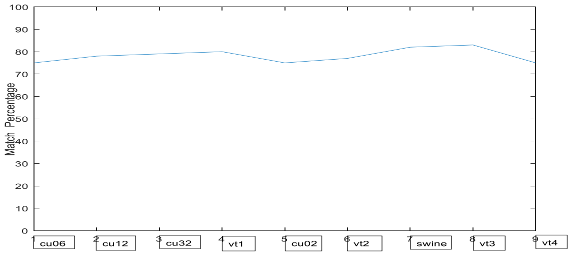


Fig.15 Plot showing Match Point vs VT signal selected

RESULTS & DISCUSSIONS

In this comparative study different feature extraction algorithms for ECG waves are discussed.Pan Tompkins and DWT-MMA algorithm werestudied to extract the fiducial points. Both algorithms was found to detect the QRS width accurately when the signal was a Sinus Rhythm. But the scenario of VT and VF in both Pan-Tompkins and DWT-MMA could not be distinguished accurately. The TCI algorithm is a time-domain analysis method which utilizes a statistical approach for identifying VT/VF. The drawback of this method is that the overlapping region of VT and normal rhythm is very high leading to decision making difficult.

Today almost 50 years after the conception of the ICD and 35 years after the first human implant, ICD therapy is the treatment of choice for patients at risk for life-threatening arrhythmias either as secondary or primary prevention. Future scope of the paper lies in modifying the studied algorithms where the flaws were identified and testing it with a wider database.

REFERENCES

[1] https://www.ahajournals.org/doi/abs/10.1161/circulationaha.114.008729
[2] Implantable Cardioverter-Defibrillator A Practical Manual L.BingLiem 2001
[3] J.an andW. J. Tompkins, “A real-time QRS detection algorithm,” IEEE Trans. Biomed. *Eng.*, vol. 32, no. 3, pp. 230–236, Mar. 1985.
[4] Qibin Zhao, and Liqing Zhan, “ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines,” International Conference on Neural Networks and Brain, ICNN&B ’05, vol. 2, pp. 1089-1092, 2005
[5] ECG Workout Exercises in arrhythmia interpretation Jane Huff, 6ht edition.
[6] https://in.mathworks.com/help/matlab/creating_guis/about-the-simple-guide-gui-example.html
[7] B. Castro, D. Kogan, and A. B. Geva, “ECG feature extraction using optimal mother wavelet,” The 21st IEEE Convention of the Electrical and Electronic Engineers in Israel, pp. 346-350, 2000.
[8] V. S. Chouhan, and S. S. Mehta, “Detection of QRS Complexes in 12-lead ECG using Adaptive Quantized Threshold,” IJCSNS International Journal of Computer Science and Network Security, vol. 8, no. 1, 2008.
[9] Evangelos, Dwaipayan Biswas, Amit Acharyya, Taihai Chen, KoushikMaharatna, James Rosengarten, John Morgan and Nick Curzen , “A Low-Complexity ECG Feature Extraction Algorithm for Mobile Healthcare Applications” , IEEE Journal of Biomedical and Health Informatics, Vol. 17, No. 2, March 2013.
[10] NareshVemishetty, PravanjanPatra ,“Low Power Personalized ECG based System Design Methodology for remote cardiac health monitoring” , IEEE Access (Volume: 4) , **pp.** 8407 – 8417, **Electronic ISSN:** 2169-3536, 29 November 2016,
[11] S.Karpagachelvi, Dr.M.Arthanari and M.Sivakumar, “ECG Feature Extraction Techniques - A Survey Approach,” (IJCSIS) International Journal of Computer Science and Information Security,Vol. 8, No. 1, April 2010
[12] https://www.instructables.com/id/Stereo-Audio-with-Arduino/
[13] ”Ventricular Tachycardia and Fibrillation Detection by a Sequential Hypothesis Testing Algorithm”,Nitish.V.Thakor,Yi-Sheng ZHU, IEEE Transitions on Biomedical Engineering,vol.37,N0.9,September 1990.
[14] Zheng XS, Zhu Ys ,Thakor NV , Wang Zz , “ Detecting Ventricular Tachycardia and Fibrillation by complexity measure,” IEEE Transaction on biomedical engineering 1999, 46(5):548-555
[15] https://physionet.org/cgi-bin/atm/ATM
[16] ML Brown, S Sarkar, “Discrimination of Shockable Tachycardia by Morphologic Stability of Ventricular Electrogram” IEEE Conferences on Computers in Cardiology 2005,32:191-194.
[17] https://physionet.org/
[18] Tom Kenny, ‘The Nexiostence of the uts and Bolts of ICD Therapy” 2006
[19] Roland X. Stroobandt, S. Serge Barold ,Alfons F. Sinaeve, “Implantable Cardioverter-Defibrillators Step by Step